Nonlinear entanglement witnesses, covariance matrices and the geometry of separable states

Otfried Gühne

Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, A-6020 Innsbruck, Austria

E-mail: `otfried.guehne@uibk.ac.at`

Norbert Lütkenhaus

Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

E-mail: `nlutkenhaus@iqc.ca`

Abstract. Entanglement witnesses provide a standard tool for the analysis of entanglement in experiments. We investigate possible nonlinear entanglement witnesses from several perspectives. First, we demonstrate that they can be used to show that the set of separable states has no facets. Second, we give a new derivation of nonlinear witnesses based on covariance matrices. Finally, we investigate extensions to the multipartite case.

1. Introduction

Entanglement plays an outstanding role in many protocols of quantum information theory. Consequently it is under intensive research from several perspectives: from the theoretical side many efforts are undertaken to recognize it via separability criteria [1, 2, 3, 4, 5] or to quantify it via entanglement measures [6]. From the experimental side, a huge amount of work is devoted to the experimental generation of entanglement using photons [7, 8, 9], ions in a trap [10, 11], or solid state systems [12].

To confirm the success of such an experiment, one has to verify that entanglement was indeed produced. Here it is important to perform the analysis without making use of hidden assumptions concerning the state [13]. Entanglement witnesses are a versatile tool for this entanglement verification [2, 7, 9, 14, 15, 16]. They are observables which have, by construction, a positive expectation value on all separable states, hence a negative expectation value signals the presence of entanglement. Besides the mere detection, they also allow for a quantitative analysis by giving bounds on entanglement measures [17, 18]. Consequently, they are now widely used in experiments.

In Ref. [19] it has been shown that one can improve all entanglement witnesses for bipartite systems by nonlinear correction terms. In this paper we extend the analysis of these nonlinear entanglement witnesses in several directions. First, in Section 2, we recall some facts concerning entanglement and entanglement witnesses. We explain the underlying definitions, their geometrical interpretation and the main idea for nonlinear witnesses. In Section 3 we show
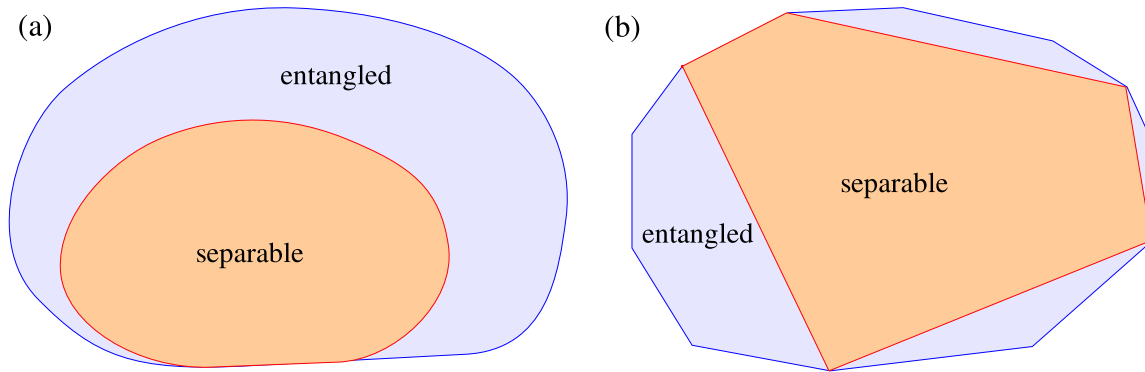
Figure 1. (a) Schematic view of the set of separable states as a convex subset of the convex set of all states. (b) A different view on the same fact. The set of states as a polytope with the pure states as corners. Some of these corners correspond to product states and they span the set of separable states. Note, however, that in reality there are infinitely many pure states.

how nonlinear witnesses can be constructed starting from any witness for bipartite systems. We follow the proof from Ref. [19] here and use it to derive that the set of separable states has no facets. In Section 4 we give an alternative proof for the fact that any entanglement witness can be improved. This proof uses covariance matrices of special observables for the derivation, highlighting the close connection between the theory of nonlinear witnesses and separability criteria in terms of covariance matrices [5, 20, 21, 22]. Finally, in Section 5 we discuss to which extent the presented methods may be used to derive nonlinear entanglement witnesses for the multipartite case.

2. Separability and the idea of nonlinear witnesses

Let us first recall the definition of entanglement and separability [23]. By definition, a quantum state ϱ shared between Alice and Bob is separable if it can be written as a mixture of product states, that is

$$\varrho = \sum_i p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i|, \tag{1}$$

where $p_i \geq 0$ and $\sum_i p_i = 1$, that is, the p_i form a probability distribution. If this is not the case, then ϱ is called entangled.

Physically, this definition means that a separable state can be produced using local operations and classical communication: by public communication, Alice and Bob can agree on producing the states $|a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i|$ locally and agree on the probabilities p_i for these states.

Geometrically, this definition implies that the set of separable states is a convex set. First, the set of all states, that is, all positive matrices with $Tr(\varrho) = 1$ is a convex set. Then the set of separable states is a convex subset. It has the product states as its extremal points, since any separable state can be written as a convex combination of product states. Two possible schematic views of this situation are shown in Fig. 1. However, one has to be careful with such schematic pictures, since these considered convex sets are high-dimensional and a two-dimensional plot can never characterize all features. But, as we will see in this paper, one can derive some general statements on the geometry of these sets.

The question, whether a given state is separable or entangled is the so called separability problem. In order to answer this question, many separability criteria have been proposed, but none of them delivers a complete solution of the problem. A famous criterion is the criterion of the positivity of the partial transposition (PPT criterion, [1]). It is defined as follows: given a
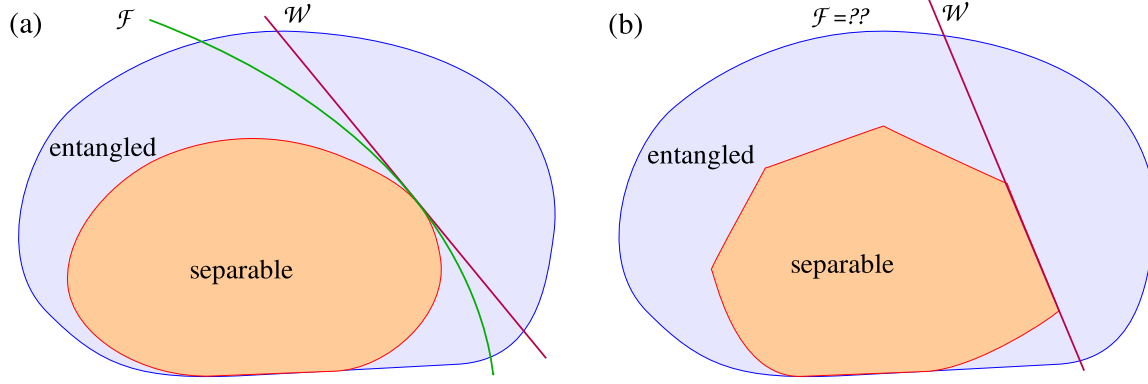
Figure 2. (a) Schematic view of an entanglement witness. The line $\mathcal{W}$ denotes the states where $Tr(\varrho\mathcal{W}) = 0$. The curve $\mathcal{F}$ sketches a possible nonlinear witness. (b) If the set of separable states has some facets, it is not clear how to find a nonlinear witness. See discussion in the text.

quantum state ϱ in a product basis

$$\varrho = \sum_{ij,kl} \varrho_{ij,kl}|i\rangle\langle j| \otimes |k\rangle\langle l|, \tag{2}$$

its partial transpose with respect to Bob's system is defined as

$$\varrho^{T_B} = \sum_{ij,kl} \varrho_{ij,lk}|i\rangle\langle j| \otimes |k\rangle\langle l|. \tag{3}$$

If ϱ is a separable state and has a representation as in Eq. (1) it can be easily seen that the partial transpose is a valid state and hence a matrix with only positive eigenvalues, i.e., $\varrho^{T_B} \geq 0$. Thus, if for a state the partial transpose is not positive (ϱ is NPT), the state must be entangled. Indeed, it has been shown [2] that for 2×2 and 2×3 systems a state is PPT if and only if it is separable, while for other dimensions there are also PPT entangled states. Besides the PPT criterion there are, however, many other criteria, which may detect a state if the PPT criterion fails [3, 4, 5].

A different approach to the separability problem uses entanglement witnesses [2, 14, 15, 16]. As already mentioned, these are observables with a positive mean value on all separable states, so a negative expectation value guarantees that the state is entangled. Geometrically, witnesses can be seen as hyperplanes which separate some entangled states from the set of separable states (see Fig. 2 (a)). Here, the hyperplane, indicated as a line corresponds to the states with $\langle\mathcal{W}\rangle = Tr(\mathcal{W}\varrho) = 0$.

From the fact that the set of separable states is convex, it can easily be deduced that for any entangled state there exist a witness detecting it. Finding such a witness, however, is not easy, since, as already mentioned, the separability problem is not solved yet. But if a state violates a certain separability criterion witnesses can often be directly constructed.

To give an example, for any NPT state ϱ_0 we find that $\varrho_0^{T_B}$ has a negative eigenvalue λ_- and a corresponding eigenvector $|\phi\rangle$. Now, an entanglement witness for this state is given by

$$\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}. \tag{4}$$

Indeed, due to the identity $Tr(XY^{T_B}) = Tr(X^{T_B}Y)$ which holds for arbitrary matrices X, Y we have $Tr(\varrho_0\mathcal{W}) = Tr(\varrho_0^{T_B}|\phi\rangle\langle\phi|) = \lambda_- < 0$ while for separable (and hence PPT) states we have

$Tr(\varrho\mathcal{W}) = \langle\phi|\varrho^{T_B}|\phi\rangle \geq 0$. For the following discussion it is important to note that the witness in Eq. (4) is not specific: since the PPT criterion is necessary and sufficient for low dimensions, witnesses of this type suffice to detect all states in these systems. Furthermore, such witnesses can be shown to be optimal [15], i.e. there is no linear witnesses which detects the same states as $\mathcal{W}$ and some states in addition.

In view of Fig. 2(a) it is now a natural question to ask whether one can improve a linear witness by some nonlinear functional. Generically, a witness gives rise to a linear functional $\mathcal{F}_l(\varrho) = Tr(\mathcal{W}\varrho)$ and a state ϱ is detected whenever $\mathcal{F}_l(\varrho) < 0$. The aim would be to find a nonlinear functional $\mathcal{F}_{nl}$ of the type

$$\mathcal{F}_{nl}(\varrho) = Tr(\mathcal{W}\varrho) - \mathcal{X}(\varrho), \tag{5}$$

which still should be positive on all separable states. Since we are looking for experimentally implementable entanglement conditions, the nonlinearity $\mathcal{X}(\varrho)$ should be a function of some expectation values of observables. As we will see, one can take a quadratic polynomial of certain expectation values. It is reasonable to consider only $\mathcal{F}_{nl}(\varrho)$ which are stronger than the linear $\mathcal{F}_l(\varrho)$. So we require that $\mathcal{F}_{nl}(\varrho)$ detects all the states that are detected by $\mathcal{F}_l(\varrho)$, and some states in addition.

However, in view of Fig. 2(b) is is not so clear, that any witness can be improved by some quadratic terms. Indeed, it might be that certain surfaces of the set of separable states are not curved, and hence some witnesses cannot be improved in the way described above. To investigate this phenomenon, we need some more terminology [24].

For a given observable A we call the set $H_{A,a} = \{\varrho : Tr(A\varrho) \leq a\}$ a half-space. For instance, the states detected by a witness $\mathcal{W}$ form just the half-space $H_{\mathcal{W},0}$. A boundary $\pi_{A,a} = \{\varrho : Tr(A\varrho) = a\}$ is called a hyperplane. Further, let F be subset of a compact convex set D in a d-dimensional space. We call F a face, if there exists a half-space $H_{A,a}$ with $D \subseteq H_{A,a}$ and $F = D \cap \pi_{A,a}$ for some A and a. Note that in this definition $H_{A,a}$ is usually not unique. Geometrically, a face is just a set of points at the border of D, where all points lie in some hyperplane. If a face is of the maximal dimension $d - 1$, we call the face a *facet*. Then the half-space $H_{A,a}$ is unique.

Concerning separability, it has been shown in Ref. [24] that the set of separable states is not a polytope. That is, for a description of its borders it is not sufficient to consider a *finite* number of hyperplanes. One can also say the the border does not consist of facets only, hence it must be curved at some points. The connection between facets and nonlinear witnesses becomes also clear: if the set of separable states has a facet F as in Fig. 2(b), then this facet corresponds to unique $H_{A,a}$. In other words, it defines a witness $\mathcal{W}$. This witness now can not be improved by a correction like

$$\mathcal{F}(\varrho) = \langle\mathcal{W}\rangle - \langle X\rangle^2 \tag{6}$$

for some hermitian X : Since F has dimension $d - 1$ and $\mathcal{F}(\varrho) = 0$ for all $\varrho \in F$ this implies that $X = \alpha\mathcal{W}$ for some real α. Hence, $\mathcal{F}(\varrho)$ does not detect any state in addition to $\mathcal{W}$. So the existence of facets on the set of separable states is closely related to the existence of nonlinear witnesses.

3. Basic results on nonlinear witnesses and their geometrical interpretation

Let us now explain the method of Ref. [19] for the construction of nonlinear witnesses. We consider first the witness from Eq. (4).

As a starting point note that a functional like $F(\varrho) = \langle X\rangle\langle X^\dagger\rangle$ for any X is convex in the state. Convexity means that if $\varrho = \sum_k p_k\varrho_k$ is a convex combination of some states, then $F(\varrho) \leq \sum_k p_k F(\varrho_k)$. This fact can be directly calculated, see. e.g. [25]. Consequently, it implies

that a functional like $\mathcal{G} = \langle A \rangle - \sum_i \alpha_i \langle X_i \rangle \langle X_i^\dagger \rangle$ with $\alpha_i \geq 0$ is concave in the state. So for convex combinations we have $\mathcal{G}(\varrho) \geq \sum_k p_k \mathcal{G}(\varrho_k)$.

Let us assume that we have taken $A = |\phi\rangle\langle\phi|, X_i = |\phi\rangle\langle\psi_i|$ for an arbitrary $|\psi_i\rangle$ and take a separable state ϱ. Then the partial transpose of the state ϱ^{T_B} is again separable and can be written as a convex combination of product states, $\varrho^{T_B} = \sum_k p_k |a_k b_k\rangle\langle a_k b_k|$. For a product vector $|\chi\rangle = |a_k b_k\rangle$ we have

$$\mathcal{G}(|\chi\rangle\langle\chi|) = \langle\chi|\phi\rangle\langle\phi|\chi\rangle \cdot [1 - \sum_i \alpha_i \langle\chi|\psi_i\rangle\langle\psi_i|\chi\rangle]$$

$$=: \langle\chi|\phi\rangle\langle\phi|\chi\rangle \cdot P(\chi) \tag{7}$$

Thus, if the function $P(\chi)$ is positive on all product states, $\mathcal{G}$ is positive on all product states. Then, by concavity, it is also positive on convex combinations thereof, hence $\mathcal{G}$ is positive on all separable ϱ^{T_B}. Consequently, with the chosen X_i,

$$\mathcal{F}(\varrho) = \mathcal{G}(\varrho^{T_B}) = \langle|\phi\rangle\langle\phi|^{T_B}\rangle - \sum_i \alpha_i \langle X_i^{T_B}\rangle\langle (X_i^{T_B})^\dagger\rangle \tag{8}$$

is positive on all separable states and is hence a nonlinear improvement of the witness $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$. From this we have:

Theorem 1. *(a) Let $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ be an entanglement witness. We define $X_i = |\phi\rangle\langle\psi|$ for an arbitrary $|\psi\rangle$ and choose $s(\psi)$ as the square of the largest Schmidt coefficient of $|\psi\rangle$. Then*

$$\mathcal{F}^{(1)}(\varrho) = \langle|\phi\rangle\langle\phi|^{T_B}\rangle - \frac{1}{s(\psi)}\langle X^{T_B}\rangle\langle (X^{T_B})^\dagger\rangle \tag{9}$$

is a nonlinear improvement of $\mathcal{W}$.
(b) If we take the same $\mathcal{W}$ and define $X_i = |\phi\rangle\langle\psi_i|, \quad i = 1, ..., K$ with some orthonormal basis $|\psi_i\rangle$, then

$$\mathcal{F}^{(2)}(\varrho) = \langle|\phi\rangle\langle\phi|^{T_B}\rangle - \sum_{i=1}^{K} \langle X_i^{T_B}\rangle\langle (X_i^{T_B})^\dagger\rangle \tag{10}$$

is also a nonlinear witness which improves $\mathcal{W}$.

Proof. (a) In order to show this, note that the squared overlap between a state $|\psi\rangle$ and a product state is bounded by the maximal squared Schmidt coefficient [7], that is, $|\langle\psi|\chi\rangle|^2 \leq s(\psi)$. From this it directly follows that $P(\chi)$ in Eq. (7) is positive. (b) For the $|\psi_i\rangle$ we have in Eq. (7) $\sum_i \langle\chi|\psi_i\rangle\langle\psi_i|\chi\rangle = Tr(|\chi\rangle\langle\chi|) = 1$ which implies that $P(\chi) = 0$ and proves the claim. $\quad\square$

To give an example how this looks like, let us consider the two-qubit case. We take $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ with $|\phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. This witness can be written as

$$\mathcal{W} = \frac{1}{4}(\mathbb{1} \otimes \mathbb{1} + \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z). \tag{11}$$

This representation shows that $\mathcal{W}$ can be evaluated by measuring $\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y$ and $\sigma_z \otimes \sigma_z$, and it can be shown that these three measurements are the optimal ones [16]. To improve the witness, we take $|\psi\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, then a direct calculation using Theorem 1(a) leads to the nonlinear witness

$$\mathcal{F}^{(1)}(\varrho) = \langle\mathcal{W}\rangle - \frac{1}{8}\langle\sigma_x \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_x\rangle^2 - \frac{1}{8}\langle\sigma_y \otimes \sigma_z - \sigma_z \otimes \sigma_y\rangle^2. \tag{12}$$

If we consider Theorem 1(b) and take the $|\psi_i\rangle$ as the four Bell states, we arrive at

$$\mathcal{F}^{(2)}(\varrho) = \langle\mathcal{W}\rangle - \frac{1}{16}\Big(\langle\sigma_x \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_x\rangle^2 + \langle\sigma_y \otimes \sigma_z - \sigma_z \otimes \sigma_y\rangle^2 + \langle\sigma_y \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_y\rangle^2$$

$$+ \langle\sigma_x \otimes \sigma_z - \sigma_z \otimes \sigma_x\rangle^2 + \langle\sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z\rangle^2 + \langle\sigma_x \otimes \sigma_y - \sigma_y \otimes \sigma_x\rangle^2 + \langle\mathcal{W}\rangle^2\Big). \tag{13}$$

Interestingly, the values of some of the quadratic terms can be determined already from the measurements like $\sigma_x \otimes \sigma_x$ etc. which were already needed to evaluate $\mathcal{W}$. Hence, the nonlinear witness can be used to improve the entanglement detection from the same data given. One should also mention that the structure of the nonlinear improvements as a sum of squares is generic: we can write the term $X^{T_B} = H + i \cdot A$ as a sum of its hermitian and anti-hermitian part, where H and A are hermitian. Then we have $\langle X^{T_B} \rangle \langle (X^{T_B})^\dagger \rangle = \langle H \rangle^2 + \langle A \rangle^2$, which leads to this structure.

Note that Theorem 1 provides a whole class of nonlinear improvements, since one can pick an arbitrary $|\psi\rangle$ and compute the corresponding nonlinearity. This freedom may be used to design nonlinear witnesses for special experimental purposes. Concerning the strength of the nonlinear improvements for the case of two qubits it has been shown in Ref. [19] that nonlinear witnesses like Eq. (12, 13) improve the witness $\mathcal{W}$ quite significantly. For the general case, we state a result from Ref. [19] without the proof:

Theorem 2. *(a) Let $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ be a witness. A state ϱ can be detected by a witness of the type $\mathcal{F}^{(1)}$ from Eq. (9) if and only if*

$$\langle\phi|\varrho^{T_B}|\phi\rangle < \left[Tr_B \left(\sqrt{Tr_A(\varrho^{T_B}|\phi\rangle\langle\phi|\varrho^{T_B})} \right) \right]^2. \tag{14}$$

(b) In the same situation, a state ϱ can be detected by a witness of the type $\mathcal{F}^{(2)}$ from Eq. (10) if and only if

$$\langle\phi|\varrho^{T_B}|\phi\rangle < \langle\phi|(\varrho^{T_B})^2|\phi\rangle \tag{15}$$

holds. In this case, the state is detected by all nonlinear witnesses of the type $\mathcal{F}^{(2)}$.
(c) Finally, if Eq. (15) is fulfilled, then also Eq. (14) holds, thus the witnesses of the type $\mathcal{F}^{(1)}$ are stronger. Furthermore, Eqs. (14, 15) are never fulfilled for PPT states.

One interesting point in this Theorem is the fact that the nonlinear improvements do not detect PPT states. The witness $\mathcal{W}$ is derived from the separability criterion of the positivity of the partial transpose and the nonlinear improvements of $\mathcal{W}$ are not more powerful than the original PPT criterion. This may sound disappointing at first sight. One should note, however, that the proof relies on special results for the PPT criterion and it is unlikely that the same fact holds also for witnesses derived from other entanglement criteria.

Let us now discuss nonlinear improvements for other entanglement witnesses, which are not related to the PPT criterion. This can be done via the theory of positive maps. Let us shortly explain this subject. Let $\mathcal{H}_B$ and $\mathcal{H}_C$ be Hilbert spaces and let $\mathcal{B}(\mathcal{H}_i)$ denote the linear operators on it. A linear map $\Lambda : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_C)$ is called positive if (a) it maps hermitian operators onto hermitian operators, fulfilling $\Lambda(X^\dagger) = \Lambda(X)^\dagger$ and (b) it preserves the positivity, i.e. if $X \geq 0$ then $\Lambda(X) \geq 0$. Note that the second condition implies that it maps valid density matrices onto density matrices (up to a normalization). A positive map Λ is called *completely* positive when for an arbitrary $\mathcal{H}_A$ the map $\mathbb{I}_A \otimes \Lambda$ is positive, otherwise, it Λ is positive, but not completely positive. Here, $\mathbb{I}_A$ denotes the identity on $\mathcal{B}(\mathcal{H}_A)$. For example, the transposition is positive, but not completely positive: while $X \geq 0$ implies $X^T \geq 0$, the *partial* transposition does not preserve the positivity of a state.

Thus, similarly as the PPT criterion, other entanglement criteria can be formulated from other positive, but not completely positive maps. Indeed, it has been shown [2, 26] that a state $\varrho \in \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ is separable if and only if for all positive maps Λ the relation

$$\mathbb{I}_A \otimes \Lambda(\varrho) \geq 0 \tag{16}$$

holds. Consequently, if ϱ is entangled there must be a positive, but not completely positive map Λ where $\mathbb{I}_A \otimes \Lambda(\varrho)$ has a negative eigenvalue λ_- and a corresponding eigenvector $|\phi\rangle$. Taking

$(\mathbb{I}_A \otimes \Lambda)^+$ as the adjoint of the map $(\mathbb{I}_A \otimes \Lambda)$ with respect to the scalar product $\langle X|Y \rangle = Tr(X^\dagger Y)$ a witness detecting ϱ is given by

$$\mathcal{W} = (\mathbb{I}_A \otimes \Lambda)^+(|\phi\rangle\langle\phi|), \tag{17}$$

since we have $Tr[\rho\mathcal{W}] = Tr[\rho(\mathbb{I}_A \otimes \Lambda)^+(|\phi\rangle\langle\phi|)] = Tr[\mathbb{I}_A \otimes \Lambda(\varrho)|\phi\rangle\langle\phi|] = \lambda_-$. By some rescaling we can always achieve that $(\mathbb{I}_A \otimes \Lambda)^+$ is not trace increasing. Then this witness can be improved as shown in Theorem 1.

Starting with an arbitrary witness, we make use of the Jamiołkowski isomorphism [26, 27] between operators and maps. According to this, an operator E on $\mathcal{B}(\mathcal{H}_B) \otimes \mathcal{B}(\mathcal{H}_C)$ corresponds to a map $\varepsilon : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_C)$ acting as

$$\varepsilon(\varrho) = Tr_B(E\varrho^T \otimes \mathbb{1}_C). \tag{18}$$

Conversely, we have

$$E = (\mathbb{I}_{B'} \otimes \varepsilon)(|\phi^+\rangle\langle\phi^+|), \tag{19}$$

where $\mathcal{H}_{B'} \cong \mathcal{H}_B$ and $|\phi^+\rangle = \sum |ii\rangle$ is a maximally entangled state on $\mathcal{H}_{B'} \otimes \mathcal{H}_B$. The important fact is that if E is an entanglement witness, then ε is a positive, but not completely positive map [26]. Again, by rescaling the witness we can achieve that the positive map is not trace increasing. Hence, any witness can be written as in Eq. (17) for a suitable positive map and we arrive at:

Theorem 3. *Any bipartite entanglement witness can be improved by nonlinear corrections. This can be done by calculating the corresponding positive map from the Jamiołkowski isomorphism and then applying the methods of Theorem 1.*

According to our discussion in the previous Section, we can directly conclude:

Theorem 4. *The set of separable states has no facets.*

The generic construction of nonlinear witnesses allows to conclude that at the border between separable and entangled states there is no facet. But also at the border between the separable states and the non-positive matrices there are no facets: any facet of this kind would correspond to a $\mathcal{W}$ which is positive. Then, in Eq. (18) the map ε is completely positive, but nevertheless one can derive from it a nonlinear functional which is positive on all separable states as in Theorem 1.

Note that the Theorem 4 does not imply that the surface of the set of separable states is a manifold which is differentiable in every point. It still may have some edges or faces, however, these edges do not have the maximal possible dimension.

4. An alternative derivation

In this Section, we give an alternative way of deriving nonlinear improvements for a given linear witness. This proof uses covariance matrices for the construction. Although the derivation is completely different, the resulting nonlinear witnesses are similar to the constructions of the previous section, they are, however, slightly weaker. The method presented here is closely related to the results of Ref. [20, 21, 22], and may be used to improve some separability criteria given in these references.

To start, let A_k be a basis of the operator space for Alice, and let B_k be a basis of the operator space for Bob. That is, for two qubits the A_k, B_k may be the Pauli matrices including the identity. For a given state ϱ we define a hermitian matrix η with the entries $\eta_{i;j} = \eta_{i_1,i_2;j_1,j_2}$ via

$$\eta(\varrho, A_k, B_k) \equiv \eta_{i_1,i_2;j_1,j_2} := \langle A_{i_1}A_{j_1} \otimes B_{i_2}B_{j_2} \rangle. \tag{20}$$

The notion $\eta(\varrho, A_k, B_k)$ should emphasize the dependence of η on the state and the observables, and $\eta_{i_1,i_2;j_1,j_2}$ emphasizes the entries of η. We mix both notations when there is no risk of

confusion. We then define the partial transposition of η as

$$\eta^{T_B} = \eta_{i_1,j_2;j_1,i_2} := \langle A_{i_1} A_{j_1} \otimes B_{j_2} B_{i_2} \rangle. \tag{21}$$

Then we have the following Lemma.

Lemma 5. *(a) We have always $\eta \geq 0$, i.e., η is a positive matrix.*
(b) The partial transposition fulfills

$$\eta^{T_B}(\varrho, A_k, B_k) = \eta(\varrho^{T_B}, A_k, B_k^T) \tag{22}$$

(c) For a state ϱ we have $\eta^{T_B} \geq 0$ if and only if ϱ is PPT.
    Proof. (a) It is known that the (asymmetric) covariance matrix

$$\gamma = \langle A_{i_1} A_{j_1} \otimes B_{i_2} B_{j_2} \rangle - \langle A_{i_1} \otimes B_{i_2} \rangle \langle A_{j_1} \otimes B_{j_2} \rangle \tag{23}$$

is always positive semidefinite [28]. The nonlinear part $\langle A_{i_1} \otimes B_{i_2} \rangle \langle A_{j_1} \otimes B_{j_2} \rangle$ of it is also positive (and subtracted), thus the linear part, which corresponds to η, must be positive. (b) This can be simply calculated, using the general fact that $Tr(X^{T_B} Y) = Tr(XY^{T_B})$. (c) The direction "$\Leftarrow$" follows already from (a) and (b). To see the other direction, assume that $\varrho^{T_B} \ngeq 0$. Then, ϱ^{T_B} must have some negative eigenvalue λ_- and a corresponding eigenvector $|\phi\rangle$. Now we can expand the operator $|\phi\rangle\langle\phi| = |\phi\rangle\langle\phi|^2 = \sum_{i_1,j_1} \alpha_{i_1,i_2} A_{i_1} \otimes B_{i_2}^T$ in the operator basis. Then we have

$$\lambda_- = Tr(\varrho^{T_B} |\phi\rangle\langle\phi|) = \sum_{i_1,i_2} \sum_{j_1,j_2} \alpha_{i_1,i_2} \eta^{T_B}(\varrho, A_k, B_k)_{i_1,i_2;j_1,j_2} \alpha_{j_1,j_2} = \langle\alpha|\eta^{T_B}|\alpha\rangle < 0, \tag{24}$$

as can be checked by direct calculation. This proves the claim. $\qquad\qquad\square$
    Before we can improve witnesses, we need one more definition. We define:

$$\chi \quad := \quad \langle A_{i_1} \otimes B_{i_2} \rangle \langle A_{j_1} \otimes B_{j_2} \rangle, \tag{25}$$
$$\Gamma \quad := \quad \eta^{T_B} - \chi. \tag{26}$$

χ is just the nonlinear part of the covariance matrix. Γ is similar to the covariance matrix, but in general it is not a covariance matrix. However, if ϱ is PPT, then Γ is the covariance matrix for the observables $A_{i_1} \otimes B_{i_2}^T$ in the state ϱ^{T_B}. Thus, in this case it is also positive.
    Let us assume that we have an NPT state ϱ and consider a witness $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ as in Eq. (4). More generally, we could also consider a witness of the type P^{T_B} where P is a positive operator. If we find a positive operator $Q = Q_{i_1,i_2;j_1,j_2} \geq 0$ such that

$$|\phi\rangle\langle\phi| = \sum_{i_1,i_2} \sum_{j_1,j_2} Q_{i_1,i_2;j_1,j_2} A_{i_1} A_{j_1} \otimes B_{i_2}^T B_{j_2}^T, \tag{27}$$

we have

$$Tr(\Gamma Q) \geq 0 \tag{28}$$

for all PPT states. However, we have also

$$Tr(\Gamma Q) = |\phi\rangle\langle\phi|^{T_B} - Tr(\chi Q). \tag{29}$$

This implies that $Tr(\Gamma Q)$ is the desired nonlinear functional which improves the witness $\mathcal{W}$.
    The question remains, whether such a Q can always be found. Indeed this is the case. We can always construct it as above in the proof of the Lemma 5 (c). This construction, however, is not very useful, since it leads to nonlinear functionals of the type $\mathcal{F} = \langle\mathcal{W}\rangle - \langle\mathcal{W}\rangle^2$ which are

not better than the witness. But we can choose other Q, since the observables $A_{i_1}A_{j_1} \otimes B_{i_2}^T B_{j_2}^T$ are tomographically overcomplete and thus Q is by no means unique. The characterization of the possible Q can be summarized as follows:

Theorem 6. *(a) The set of possible Q is closed and convex.*
(b) For entanglement detection it suffices to consider the extremal points of this set. If Q is of rank one, it is extremal.
(c) All pure extremal points can be found as follows: For the given $P \geq 0$ (e.g. $P = |\phi\rangle\langle\phi|$) one considers the spectral decomposition of $\sqrt{P}$, that is $\sqrt{P} = UDU^\dagger$ with U unitary and D diagonal. Then one considers

$$X = UDV = \sum_{i_1,i_2} \alpha_{i_1 i_2} A_{i_1} \otimes B_{i_2}^T, \tag{30}$$

with arbitrary unitary V and thus complex $\alpha_{i_1 i_2}$. One extremal Q is then given by

$$Q_{i_1,i_2;j_1,j_2} = \alpha_{i_1 i_2} \alpha_{j_1 j_2}^*. \tag{31}$$

More generally, the search for an appropriate Q can be solved via the semidefinite program [29]

$$\text{minimize} \quad Tr(\Gamma Q), \tag{32}$$
$$\text{subject to} \quad Q \geq 0, \tag{33}$$
$$P = \sum_{i_1,i_2} \sum_{j_1,j_2} Q_{i_1,i_2,j_1,j_2} A_{i_1} A_{j_1} \otimes B_{i_2}^T B_{j_2}^T. \tag{34}$$

(d) The detection power of the resulting nonlinear separability criteria does not depend on the initial choice of the A_k and B_k.

Proof. (a) This is obvious. (b) It is clear, that the extremal points suffice, since the resulting entanglement conditions are linear in Q. (c) It can be straightforwardly seen that the constructed Q are valid: We have $P = \sqrt{P}\sqrt{P} = XX^\dagger$ from this it follows that Eq. (27) holds. On the other hand, any pure extremal Q must be of the type Eq. (31). This implies that we can find a corresponding X with $XX^\dagger = P$ of the desired type. Here, we use that the singular value decomposition is unique. (d) Assume that we take other observables like $\tilde{A}_k = \sum_l C_{kl} A_l$ and $\tilde{B}_k = \sum_l D_{kl} B_l$. Then the matrices C, D must be invertible and we have

$$\eta(\varrho, \tilde{A}_k, \tilde{B}_k) = C \otimes D \, \eta(\varrho, A_l, B_l) C^T \otimes D^T, \tag{35}$$
$$\Gamma(\varrho, \tilde{A}_k, \tilde{B}_k) = C \otimes D \, \Gamma(\varrho, A_l, B_l) C^T \otimes D^T. \tag{36}$$

From Eq. (30) one can read off that α transforms like $C^T \otimes D^T \alpha(\tilde{A}_k, \tilde{B}_k) = \alpha(A_l, B_l)$, this implies that $Q(\tilde{A}_k, \tilde{B}_k) = (C^T \otimes D^T)^{-1} Q(A_l, B_l)(C \otimes D)^{-1}$. This proves the claim. $\square$

Corollary 7. *Any entanglement witness of the form $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ can be improved by some quadratic corrections.*

Proof. We only have to show that the above given improvements are not all trivial. Let us assume the contrary. This would imply that for all states ϱ with $Tr(\mathcal{W}\varrho) = 0$ and for all possible Q we would have

$$\sum_{i_1,i_2} \sum_{j_1,j_2} Q_{i_1,i_2;j_1,j_2} \langle A_{i_1} \otimes B_{i_2} \rangle \langle A_{j_1} \otimes B_{j_2} \rangle = 0. \tag{37}$$

Defining $\beta_{i_1,i_2} = \langle A_{i_1} \otimes B_{i_2} \rangle$ this may be written as $\langle \beta | Q | \beta \rangle = 0$. For a $d \times d$ system the set of all density matrices is a $d^2 \times d^2 - 1$ (real) dimensional manifold. The set of states ϱ with $Tr(\mathcal{W}\varrho) = 0$ forms a $d^2 \times d^2 - 2$ dimensional affine space F. Consequently, the possible $|\beta\rangle$ arising from the $\varrho \in F$ span a $d^2 \times d^2 - 1$ dimensional subspace. Since Q is a $d^4 \times d^4$ matrix

it has d^4 eigenvectors and nonnegative eigenvalues. So if Eq. (37) were valid, then $d^4 - 1$ of the eigenvalues would equal zero, hence all Q would be of rank one, and all valid Q would be a multiple of a fixed projector. But obviously there are more than one valid α_{i_1,i_2} in Eqs. (30, 31). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

This approach gives a different view at nonlinear entanglement witnesses. As in the previous section, it can be directly extended to other positive maps besides the partial transposition. For a witness of the type $\mathcal{W} = |\phi\rangle\langle\phi|^{T_B}$ the construction in Theorem 6 would first start with a choice of a $|\psi\rangle$ to build up $X = |\psi\rangle\langle\phi|$ and finally one arrives at the nonlinear witness as in Theorem 1(a), but without the factor $1/s(\psi)$. Since this prefactor is always larger than one, the witness from Theorem 1 is stronger.

5. The multipartite setting

Finally, let us discuss shortly the extension to the multipartite case. We concentrate on the three-qubit case, since this already suffices to state the main results.

First, it is important to note that for three parties several forms of entanglement exist [30]. We call a tripartite pure state fully separable if it is of the form $|\psi\rangle = |a\rangle \otimes |b\rangle \otimes |c\rangle$, and a mixed state is fully separable, if it is of the form

$$\varrho = \sum_i p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i| \otimes |c_i\rangle\langle c_i|. \tag{38}$$

Furthermore, a pure state is biseparable, if it is separable with respect to one of the three possible bipartitions $A|BC$, $AB|C$ or $AC|B$, e.g. $|\psi\rangle = |\phi\rangle_{AB} \otimes |\chi\rangle_C$. Again, a mixed state is called biseparable, if it can be written as a convex sum of biseparable states,

$$\varrho = \sum_i p_i |\psi_i^{(\mathrm{bs})}\rangle\langle\psi_i^{(\mathrm{bs})}|, \tag{39}$$

where the biseparable states $|\psi_i^{(\mathrm{bs})}\rangle$ might be biseparable with respect to different partitions. Otherwise, the state ϱ is called genuine multipartite entangled.

For these types of entanglement one can define entanglement witnesses as for the bipartite case. However, it is important to note that in experiments mainly genuine multipartite entanglement is of interest: to confirm the success of an experiment with three qubits one has to show that all three qubits were entangled, and not only two of them. Hence it is not sufficient to exclude full separability.

The question arises, whether we can construct nonlinear entanglement witnesses also for the multipartite case. In one case this can be done. Let $\mathcal{W}$ by a witness ruling full separability, that is, $Tr(\varrho\mathcal{W}) \geq 0$ for all fully separable states. Then we can pick the bipartition $A|BC$ and find a nonlinear improvement $\mathcal{F}_{A|BC}(\varrho)$ for this bipartition. Then, the minimum over all bipartitions,

$$\mathcal{F}_{\mathrm{tot}}(\varrho) = \min\{\mathcal{F}_{A|BC}(\varrho), \mathcal{F}_{AB|C}(\varrho), \mathcal{F}_{AC|B}(\varrho)\}, \tag{40}$$

is clearly positive on all fully separable states. Hence, it defines a nonlinear improvement of $\mathcal{W}$.

For the more interesting case of witnesses for genuine multipartite entanglement this recipe does not work anymore. Here, the first problem comes from the fact that in the definition of biseparability in Eq. (39) also convex combinations of biseparable states with respect to different bipartitions are allowed. One might be tempted to consider in analogy to Eq. (40) a function like $\mathcal{F}_{\mathrm{tot}}(\varrho) = \max\{\mathcal{F}_{A|BC}(\varrho), \mathcal{F}_{AB|C}(\varrho), \mathcal{F}_{AC|B}(\varrho)\}$ to improve a witness for genuine multipartite entanglement. While this functional is positive on all pure biseparable states, it is however, not necessarily positive on mixed biseparable states, since the maximum of concave functions is not concave.

Another attempt for the improvement of witnesses for genuine multipartite entanglement is to calculate the $\mathcal{F}_{A|BC}(\varrho)$ etc. as above, and then one can investigate the quadratic terms for the $\mathcal{F}$ as in Eqs. (12, 13). If one finds a quadratic term which occurs in all $\mathcal{F}$, then this term may be subtracted from $\mathcal{W}$, arriving at a valid witness. However, it seems quite difficult to find a single example where this recipe works. So the search for nonlinear entanglement witnesses for multipartite systems remains an interesting and open problem for further study.

6. Conclusion

In conclusion we investigated nonlinear entanglement witnesses from different perspectives. We demonstrated that they can be used to show that the set of separable states has no facets. We also gave a new derivation of nonlinear witnesses based on covariance matrices. This highlights the close connection between nonlinear entanglement detection and separability criteria in terms of covariance matrices. Finally, we discussed the problems which occur if one wishes to construct nonlinear witnesses for the multipartite case.

Acknowledgments

References

[1] Peres A 1996 *Phys. Rev. Lett.* **77** 1413
[2] Horodecki M, Horodecki P and Horodecki R 1996 *Phys. Lett.* A **223** 1
[3] Rudolph O 2002 *Preprint* quant-ph/0202121
[4] Chen K and Wu L A 2003 *Quantum Inf. Comput.* **3** 193
[5] Gühne O, Hyllus P, Gittsovich O and Eisert J 2006 *Preprint* quant-ph/0611282
[6] Plenio M B and Virmani S 2007 *Quant. Inf. Comp.* **7** 1
[7] Bourennane M, Eibl M, Kurtsiefer C, Gaertner S, Weinfurter H, Gühne O, Hyllus P, Bruß D, Lewenstein M and Sanpera A 2004 *Phys. Rev. Lett.* **92** 087902
[8] Walther P, Resch K J, Rudolph T, Schenck E, Weinfurter H, Vedral V, Aspelmeyer M and Zeilinger A 2005 *Nature* **434** 169
[9] Lu C Y, Zhou X Q, Gühne O, Gao W B, Zhang J, Yuan Z S, Goebel A, Yang T and Pan J W 2006 *Preprint* quant-ph/0609130
[10] Leibfried D, Knill E, Seidelin S, Britton J, Blakestad R B, Chiaverini J, Hume D B, Itano W M, Jost J D, Langer C, Ozeri R, Reichle R and Wineland D J 2005 *Nature* **438** 639
[11] Häffner H, Hänsel W, Roos C F, Benhelm J, Chek-al-kar D, Chwalla M, Körber T, Rapol U D, Riebe M, Schmidt P O, Becher C, Gühne O, Dür W and Blatt R 2005 *Nature* **438** 643
[12] Steffen M, Ansmann M, Bialczak R C, Katz N, Lucero E, McDermott R, Neeley M, Weig E M, Cleland A N and Martinis J M 2006 *Science* **313** 1423
[13] van Enk S J, Lütkenhaus N and Kimble H J 2006 *Preprint* quant-ph/0611219
[14] Terhal B 2000 *Phys. Lett.* A **271** 319
[15] Lewenstein M, Kraus B, Cirac J I and Horodecki P 2000 *Phys. Rev.* A **62** 052310
[16] Gühne O, Hyllus P, Bruß D, Ekert A, Lewenstein M, Macchiavello C and Sanpera A 2002 *Phys. Rev.* A **66** 062305
[17] Gühne O, Reimpell M and Werner R F 2006 *Preprint* quant-ph/0607163
[18] Eisert J, Brandão F G S L and Audenaert K M R 2006 *Preprint* quant-ph/0607167
[19] Gühne O and Lütkenhaus N 2006 *Phys. Rev. Lett.* **96** 170502
[20] Shchukin E and Vogel W, 2005 *Phys. Rev. Lett.* **95** 230502
[21] Rigas J, Gühne O and Lütkenhaus N 2006 *Phys. Rev.* A **73** 012341
[22] Miranowicz A, Piani M, Horodecki P and Horodecki R 2006 *Preprint* quant-ph/0605001
[23] Werner R F 1989 *Phys. Rev.* A **40** 4277
[24] Ioannou L M and Travaglione B C 2006 *Phys. Rev.* A **73** 052314
[25] Hofmann H F and Takeuchi S 2003 *Phys. Rev.* A **68** 032103
[26] Horodecki M, Horodecki P and Horodecki R, 2001 *Preprint* quant-ph/0109124
[27] Jamiołkowski A 1972 *Rep. Mat. Phys.* **3** 275
[28] Robertson H P 1934 *Phys. Rev.* **46** 794

[29] Vandenberghe L and Boyd S 1996 *SIAM Review* **38** 49

[30] Acín A, Bruß D, Lewenstein M and Sanpera A 2001 *Phys. Rev. Lett.* **87** 040401